2017

The impact of the bovine faecal microbiome on *Escherichia coli* O157:H7 prevalence and enumeration in naturally infected cattle

M. Kim
USDA-ARS, moon.kim@ars.usda.gov

L. A. Kuehn
USDA-ARS, Larry.Kuehn@ars.usda.gov

J. L. Bono
USDA-ARS, jim.bono@ars.usda.gov

E. D. Berry
USDA-ARS, elaine.berry@ars.usda.gov

N. Kalchayanand
USDA-ARS, norasak.kalchayanand@ars.usda.gov

See next page for additional authors

Journal of Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

The impact of the bovine faecal microbiome on *Escherichia coli* O157:H7 prevalence and enumeration in naturally infected cattle

M. Kim[1], L.A. Kuehn[1], J.L. Bono[1], E.D. Berry[1], N. Kalchayanand[1], H.C. Freetly[1], A.K. Benson[2] and J.E. Wells[1]

1 U.S. Meat Animal Research Center, USDA, ARS, Clay Center, NE, USA
2 Department of Food Science and Technology, University of Nebraska-Lincoln, Lincoln, NE, USA

Keywords
cattle, *Escherichia coli* O157:H7, intestinal, microbiota, pathogen shedding.

Correspondence
James Wells, U.S. Meat Animal Research Center, U.S. Department of Agriculture, PO Box 166, State spur 18D, Clay Center, NE 68933, USA.
E-mail: jim.wells@ars.usda.gov

Present address: M. Kim, Animal Nutrition and Physiology Team, Department of Animal Biotechnology and Environment, National Institute of Animal Science, RDA, Wanju 55365, Korea

2017/0671: received 6 April 2017, revised 5 July 2017 and accepted 10 July 2017

doi:10.1111/jam.13545

Abstract

Aims: The objective of this study was to determine if the faecal microbiome has an association with *Escherichia coli* O157:H7 prevalence and enumeration.
Methods and Results: Pyrosequencing analysis of faecal microbiome was performed from feedlot cattle fed one of three diets: (i) 94 heifers fed low concentrate (LC) diet, (ii) 142 steers fed moderate concentrate (MC) diet, and (iii) 132 steers fed high concentrate (HC) diet. A total of 322 585 OTUs were calculated from 2,411,122 high-quality sequences obtained from 368 faecal samples. In the LC diet group, OTUs assigned to the orders Clostridiales and RF39 (placed within the class Mollicutes) were positively correlated with both *E. coli* O157:H7 prevalence and enumeration. In the MC diet group, OTUs assigned to *Prevotella copri* were positively correlated with both *E. coli* O157:H7 prevalence and enumeration, whereas OTUs assigned to *Prevotella stercorea* were negatively correlated with both *E. coli* O157:H7 prevalence and enumeration. In both the MC diet group and the HC diet group, OTUs assigned to taxa placed within Clostridiales were both positively and negatively correlated with both *E. coli* O157:H7 prevalence and enumeration. However, all correlations were weak. In both the MC diet group and the HC diet group, stepwise linear regression through backward elimination analyses indicated that these OTUs were significantly correlated ($P < 0{\cdot}001$) with prevalence or enumeration, explaining as much as 50% of variability in *E. coli* O157:H7 prevalence or enumeration.
Conclusions: Individual colonic bacterial species have little impact on *E. coli* O157:H7 shedding but collectively groups of bacteria were strongly associated with pathogen shedding.
Significance and Impact of the Study: Bacterial groups in the bovine colon may impact faecal shedding of the zoonotic pathogen *E. coli* O157:H7, and manipulation of the intestinal microbiota to alter these bacteria may reduce shedding of this pathogen and foodborne illnesses.

Introduction

Escherichia coli O157:H7 is a Shiga toxin-producing pathogen that is commonly detected in bovine faeces and can be transmitted to humans through contaminated food or water. Infection with *E. coli* O157:H7 in humans can result in illness such as bloody diarrhoea, haemorrhagic colitis, thrombotic thrombocytopenic purpura or haemolytic uraemic syndrome (Berry and Wells 2010).

In a feedlot, the majority of cattle are low-level shedders of *E. coli* O157:H7, and the presence of supershedders shedding high numbers of *E. coli* O157:H7 in faeces

($\geq 10^4$ CFU per gram) have been identified at a low incidence (Omisakin *et al.* 2003; Robinson *et al.* 2004; Brichta-Harhay *et al.* 2007). However, the supershedders of *E. coli* O157:H7 greatly contribute to *E. coli* O157:H7 contamination in an environment (Berry and Wells 2010). Faecal shedding of *E. coli* O157:H7 also was positively correlated with hide contamination of *E. coli* O157:H7 (Arthur *et al.* 2007, 2009). It has been reported that high level and continuous shedding of *E. coli* O157:H7 are related to colonization at the terminal rectum (Low *et al.* 2005). However, few studies have been conducted to find the detailed mechanisms to explain passage and colonization that subsequently lead to high-level shedding of *E. coli* O157:H7 in faeces (Berry and Wells 2010).

Recently, the high-throughput sequencing of 16S rRNA gene amplicons has provided deeper sequencing coverage to investigate the structure of bovine faecal microbiome, and results indicate that there is diversity across diets and between animals. Dowd *et al.* (2008) analysed samples from 20 dairy cattle fed a formulated standard ration for lactating animals to investigate the structure of faecal microbiome. Although the animals were fed an identical diet, only 16 of the 142 detected genera were observed across all 20 cattle in the study. Callaway *et al.* (2010) observed 18 common genera in the structure of the faecal microbiome from six cattle fed 0, 25 or 50% dried distillers grain (DG) (two cattle per diet), but only one of these genera differed across three diet groups. Shanks *et al.* (2011) used next-generation sequencing to investigate the structure of the faecal microbiome of 30 cattle equally divided across six cattle populations (five animals per population). These populations were distributed over four locations and were classified into one of three broadly defined diet groups, which consisted of a forage, a processed-grain or an unprocessed grain type of diet. A total of 633 877 cleaned sequences were obtained, and the authors reported that the faecal microbiome structure was correlated with starch concentration in the faeces of cattle. Rice *et al.* (2012) also used the next-generation sequencing technique to examine the faecal microbiome recovered from 20 cattle fed 5 diets with differing types and levels of DG, and observed the structure of the faecal microbiome was significantly different between DG-based diets and the control diet.

Gut microbiome may play an important role in controlling pathogens via direct interaction with each other and stimulation of host immunity (Kamada *et al.* 2013). Therefore, bovine faecal microbiome also may affect *E. coli* O157:H7 passage through the gastrointestinal tract and colonization thus affecting the prevalence and concentration in faeces. Recently, we analysed bovine faecal microbiome from 426 cattle fed three different diets using the high-throughput pyrosequencing method and identified 434 different genera in 21 phyla (Kim *et al.* 2014). This large study clearly indicated that bovine faecal microbiome differed by diet group, particularly between forage and grain diets. In the current study, we analysed *E. coli* O157:H7 prevalence and enumeration from the same samples and then examined the impact of faecal microbiome on *E. coli* O157:H7 prevalence and enumeration for each diet group. Understanding the variance explained by bovine faecal microbiome towards *E. coli* O157:H7 colonization and shedding using stepwise regression through backward elimination (step-down) model-building techniques (Neter *et al.* 1985) may provide new insights into strategies to understand the complex ecosystem driving faecal *E. coli* O157:H7 shedding in cattle.

Materials and methods

Animal, diets and faecal sample collection

All animal procedures were reviewed and approved by U.S. Meat Animal Research Center (USMARC) and University of Nebraska-Lincoln (UNL) Animal Care and Use Committees.

Faecal samples were collected from 426 cattle fed one of the three diets (Kim *et al.*, 2014): (i) 94 heifers fed the low maize, forage-based (low concentrate, LC) diet composed of 70% maize silage and 30% alfalfa haylage in 2010, (ii) 142 steers fed the moderate maize (MC) diet composed of 66% dry-rolled maize, 26% maize silage and 8% supplement in 2009, and (iii) 132 steers fed the high maize (HC) diet composed of 83% dry-rolled maize, 13% maize silage and 4% supplement in 2010. Each animal was restrained for <1 min in a cattle squeeze chute for faecal collections every 2 weeks between June and September of each year. Each faecal sample was collected by rectal grab using a clean gauntlet glove (NASCO, Ft. Atkinson, WI) and immediately transferred into a clean closable plastic bag for transport to the laboratory. Samples were transported to the laboratory within 20 min for analyses of *E. coli* O157:H7 (Wells *et al.* 2011). A subsample was immediately collected and frozen for later compilation, DNA extraction and 16S RNA amplicon sequence determinations as described previously (Kim *et al.*, 2014).

Determination of *E. coli* O157:H7 prevalence

A 10 g (±0·1 g) subsample of each faecal sample was placed in a sterile filter bag (NASCO) and 90 ml of sterile tryptic soy broth (Difco, BD, Sparks, MD) containing 100 mmol l^{-1} potassium phosphate buffer (18 mmol l^{-1} KH_2PO_4 and 82 mmol l^{-1} K_2HPO_4, pH 7·2; Sigma

Chemicals, St. Louis, MO) was added. Filter bags with enriched faecal samples were mixed by hand massage, and 0·5 ml was removed to a sterile tube for enumeration (described below). Samples were enriched by incubation for 2 h at 25°C followed by 6 h at 42°C, then the enriched samples were held at 4°C overnight. Enriched samples were mixed by hand massage and 0·5 ml was transferred to a 96-well deep-well block containing 0·5 ml phosphate-buffered saline with Tween (PBS-Tween; Sigma Chemicals) and 20 μl anti-O157 immunomagnetic beads (Invitrogen Corp., Carlsbad, CA). The deep-well block with enriched sample and beads were mixed for 15 min at room temperature on a vibrating VWR Incubating Microplate Shaker (VWR, Radnor, PA). The beads were removed from the sample, washed in PBS-Tween twice, and eluted in 100 μl PBS-Tween using a Kingfisher 96 robotic processor (Thermo Scientific, Waltham, MA). A 50-μl aliquot of the bead suspension was spread plated onto ntCHROMO157 (CHROMO157 Agar DRG International, Mountainside, NJ) supplemented with 5 mg of novobiocin per l (Sigma Chemicals) and 2·5 mg of potassium tellurite per l (Sigma Chemicals). The plates were incubated overnight at 37°C. Presumptive colonies were tested with DrySpot *E. coli* O157 (Oxoid Ltd, Basingstoke, UK). At least two agglutination-positive colonies were picked per plate and confirmed by PCR assay for combination of O157, H7 flagella, intimin, and Shiga toxin 1 and 2 genes (Hu *et al.* 1999).

Enumeration of *E. coli* O157:H7

To enumerate *E. coli* O157:H7 organisms in each sample, 50 μl of each faecal suspension was plated onto ntCHROMO157 agar (described above) by use of an Autoplate 4000 spiral-plater (Spiral Biotech Inc., Norwood, MA). Plates were air-dried and incubated overnight at 42°C. Presumptive colonies (up to five per sample) were tested with DrySpot *E. coli* O157 (Oxoid Ltd). Colonies with morphologies similar to the agglutination-positive colonies were counted and confirmed by PCR assay (as described above). The minimum detection limit for *E. coli* O157:H7 enumeration is 200 CFU per gram faeces. Enumeration values were log transformed for subsequent association analyses.

Pyrosequencing process

A faecal composite sample was made for each animal by pooling 0·5 g of each individual sample collected for the animal. Total community DNA was extracted from 426 faecal samples using a MiniBeadbeater-8 (BioSpec Products, Bartlesville, OK) followed by the QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany) as described previously (Martinez *et al.* 2010). The V1–V3 region of 16S rRNA genes was amplified from each DNA sample using primers 27F (5′-Adaptor primer-AGAGTTTGATCMTGGCTCAG-3′) and barcoded 518R (5′-Adaptor primer-barcode-ATTACCGCGGCTGCTGG-3′) and then was sequenced on the 454 GS FLX Titanium system as described previously (Kim *et al.* 2014).

Phylogenetic analysis

Bioinformatic programs in QIIME software package 1.6.0 (Caporaso *et al.* 2010b) were used to analyse 16S rRNA gene sequences of bovine faecal microbiome. After removing barcode and primer regions, sequences that have read lengths ≥200 bp, mean quality score ≥Q20, and homopolymer stretches <8 bp were selected for further sequence processing. The resultant sequences were aligned using the PyNAST method against the Greengenes Core reference set (DeSantis *et al.* 2006; Caporaso *et al.* 2010a) and then subjected to chimera detection using the ChimeraSlayer program (Haas *et al.* 2011). The nonchimera sequences obtained were assigned to the Greengene taxonomy against the Greengenes reference database (2012-10 release; DeSantis *et al.* 2006) using the RDP Classifier as recommended previously (Werner *et al.* 2012). Operational taxonomic units (OTUs) at 97% sequence similarity were calculated by reference-based clustering followed by de novo OTU clustering as described previously (Nelson *et al.* 2014). Proportional value for each taxon or OTU of total sequences within each diet was used to examine the impact of faecal microbiota on *E. coli* O157:H7 prevalence and enumeration. Alpha diversity indices were calculated from the 368 subsamples with 2000 sequences based on a phylogenetic tree constructed using FastTree 2.1.3 (Price *et al.* 2010).

Statistical analysis

All statistical analyses were conducted using the XLSTAT statistical software (Addinsoft, New York, NY). Proportion value of the total sequences for taxa or OTUs was log transformed. Simple linear regression was conducted to examine how *E. coli* O157:H7 prevalence or enumeration (dependent variable) varies with taxa or OTUs (explanatory variable) using the XLSTAT statistical software. Taxa or OTUs detected in <50% of total cattle in each diet group were excluded from simple linear regression. OTUs that bring significant information ($P < 0{\cdot}1$) to the model in simple linear regression were screened and then subjected to stepwise multiple linear regression with backward elimination with elimination of each least significant input until all remaining variables had a $P < 0{\cdot}05$ (Neter *et al.* 1985). Principal component

analysis (PCA) was conducted to compare faecal microbiome between the top 10% prevalence (enumeration) group and the bottom prevalence (enumeration) 10%.

Results

Data summary

Escherichia coli O157:H7 were determined by culture methods because *E. coli* O157:H7 cannot be discriminated from nonpathogenic *E. coli* using only the 16S rRNA gene partial sequence reads. Cattle fed the LC diet were sampled six times over a 10-week period. Individual animals tested for *E. coli* O157:H7 ranged one to six times positive, and on average an animal was positive 4·0 times across all LC pens. The *E. coli* O157:H7 average enumerations ranged from 1·2 to 5·1 $\log_{10}$ CFU per gram faeces and averaged 2·4 $\log_{10}$ CFU per gram faeces across all LC animals. Cattle fed the MC diet were sampled seven times over a 12-week period. Individual animals tested for *E. coli* O157:H7 ranged one to seven times positive, and on average an animal was positive 3·6 times across all MC pens. The *E. coli* O157:H7 average enumerations ranged from 1·2 to 4·9 $\log_{10}$ CFU per gram faeces and averaged 2·5 $\log_{10}$ CFU per gram faeces across all MC animals. Cattle fed the HC diet were sampled six times over a 10-week period. Individual animals tested for *E. coli* O157:H7 ranged zero to four times positive, and on average an animal was positive 1·6 times across all HC pens. The *E. coli* O157:H7 average enumerations ranged from undetectable (all samples enrichment negative for animal) to 5·1 $\log_{10}$ CFU per gram faeces and averaged 1·8 $\log_{10}$ CFU per gram faeces across all HC animals.

A total of 2 411 122 high-quality sequences were obtained from 368 samples with each sample being represented by at least 2000 sequences. The remaining 58 samples represented by <2000 sequences were excluded from bioinformatics analysis. In the LC diet group, 402 080 of the total sequences were recovered from 94 faecal samples and classified to 29 phyla. Firmicutes, candidate division TM7, Bacteroidetes, Tenericutes, Proteobacteria, Verrucomicrobia and Actinobacteria were core measurable phyla detected in ≥93 of the 94 faecal samples. Firmicutes was the first predominant phylum accounting for 69·6% of the 402 080 sequences, followed by TM7 (13·3%) and Tenericutes (6·7%).

In the MC diet group, 1 035 186 of the total sequences were recovered from 142 faecal samples and classified to 22 phyla. Firmicutes, Bacteroidetes, Proteobacteria, Tenericutes, TM7, Cyanobacteria and Actinobacteria were core measurable phyla detected in ≥137 of the 142 samples. Firmicutes was the first predominant phylum accounting for 49·0% of the 1 035 186 sequences, followed by Bacteroidetes (37·8%) and Proteobacteria (6·2%).

In the HC group, 973 856 of the total sequences were recovered from 132 faecal samples and assigned to 26 phyla. Firmicutes, Bacteroidetes, Tenericutes, Proteobacteria, Actinobacteria, Cyanobacteria and TM7 were core measurable phyla detected in ≥129 of the 132 samples. Firmicutes was the first predominant phylum and represented 77·9% of the 973 856 sequences. Bacteroidetes and Tenericutes were the second and the third predominant phyla and represented 12·0 and 4·2% of the 973 856 sequences respectively.

The 2 411 122 sequences were clustered into 322 585 OTUs at 97% sequence similarity, where 143 722 OTUs, 97 322 OTUs and 126 587 OTUs were assigned to the LC diet, the MC diet and the HC diet groups respectively. Of the 322 585 OTUs, only 42 162 OTUs were shared between at least two diet groups.

The impact of faecal microbiome on *E. coli* O157:H7 prevalence

In the LC diet group, 110 taxa and 231 OTUs were observed in 50% or more of the samples, and 7 taxa and 19 OTUs were correlated with *E. coli* O157:H7 prevalence (Tables 1 and 2) for step-down regression model building to identify relationships in phenotypic variation. Positive correlations were detected ($P < 0{\cdot}05$) for family Coprobacillaceae ($r = 0{\cdot}22$) and genus *Lactococcus* ($r = 0{\cdot}21$) placed within the phylum Firmicutes with *E. coli* O157:H7 prevalence (Table 1). On the other hand, negative correlations were detected ($P < 0{\cdot}05$) for phylum Cyanobacteria ($r = -0{\cdot}21$), and putative class 4C0d-2 ($r = -0{\cdot}29$) and putative order YS2 ($r = -0{\cdot}29$) placed within phylum Cyanobacteria. In addition, negative correlations were detected ($P < 0{\cdot}05$) for families Planococcaceae ($r = -0{\cdot}23$) and Dehalobacteriaceae ($r = -0{\cdot}26$) placed within phylum Firmicutes with *E. coli* O157:H7 prevalence.

Positive correlations were detected ($P < 0{\cdot}05$) for all 19 OTUs with *E. coli* O157:H7 prevalence (Table 2). The 19 OTUs were assigned to the putative order RF39 (four OTUs), Clostridiales (one OTU), Ruminococcaceae (seven OTUs), Catabacteriaceae (one OTU), Lachnospiraceae (one OTU), Erysipelotrichaceae (one OTU), *Ruminococcus* (two OTU), *Butyrivibrio* (one OTU) and *Bulleidia* (one OTU). No OTUs with significant negative correlations were observed. Stepwise regression with backward elimination of 33 OTUs (OTUs with $P < 0{\cdot}1$, Table S1) indicated that six OTUs had significant impact on the model ($P < 0{\cdot}05$). Multiple linear regression for six OTUs, which included two negatively associated

Table 1 Commonly distributed taxa observed in 50% or more of animals within each diet correlating (slope, nominal *P*-value, and correlation coefficient) with *Escherichia coli* O157:H7 prevalence in bovine faeces from cattle with differing levels of concentrate

Diet*	Taxonomy	Slope	*P*-value†	*R*
LC	p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Lactococcus	0·30	0·042	0·21
	p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Coprobacillaceae	0·31	0·031	0·22
	p_Cyanobacteria	−0·29	0·042	−0·21
	p_Cyanobacteria;c_4C0d-2	−0·42	0·005	−0·29
	p_Cyanobacteria;c_4C0d-2;o_YS2	−0·42	0·005	−0·29
	p_Firmicutes;c_Bacilli;o_Bacillales;f_Planococcaceae	−0·34	0·023	−0·23
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Dehalobacteriaceae	−0·37	0·011	−0·26
MC	p_Firmicutes;c_Bacilli;o_Bacillales	−0·25	0·011	−0·21
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Butyrivibrio	−0·16	0·049	−0·17
HC	p_Actinobacteria;c_Actinobacteria;o_Actinomycetales	0·16	0·042	0·18
	p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Corynebacteriaceae	0·18	0·021	0·20
	p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Corynebacteriaceae;g_Corynebacterium	0·18	0·021	0·20
	p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Propionibacteriaceae	0·18	0·013	0·22
	p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Propionibacteriaceae;g_Propionibacterium	0·16	0·023	0·20
	p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Propionibacteriaceae;g_Propionibacterium;s_acnes	0·16	0·022	0·20
	p_Firmicutes	0·02	0·029	0·19
	p_Firmicutes;c_Bacilli	0·07	0·021	0·20
	p_Firmicutes;c_Bacilli;o_Turicibacterales	0·08	0·034	0·19
	p_Firmicutes;c_Bacilli;o_Turicibacterales;f_Turicibacteraceae	0·08	0·034	0·19
	p_Firmicutes;c_Bacilli;o_Turicibacterales;f_Turicibacteraceae;g_Turicibacter	0·08	0·034	0·19
	p_Cyanobacteria	−0·10	0·032	−0·19
	p_Cyanobacteria;c_4C0d-2	−0·10	0·020	−0·20
	p_Cyanobacteria;c_4C0d-2;o_YS2	−0·11	0·015	−0·21
	p_Firmicutes;c_Erysipelotrichi	−0·09	0·014	−0·21
	p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales	−0·09	0·014	−0·21
	p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Coprobacillaceae	−0·14	0·004	−0·25

*The diets were low concentrate (LC), moderate concentrate (MC) and high concentrate (HC). Widely distributed taxa considered for each diet were 110, 137 and 146 taxa groups respectively.

†The *P*-values presented represent nominal *P*-values used for the model generation. When corrected for false discovery using Bonferroni correction, none of the individual *P*-values reached a significance threshold $P < 0{\cdot}05$ for any of the diets.

OTUs that tended to be significant when analysed individually, brought significant information to the model ($P < 0{\cdot}0001$) and explained 28·7% of variability in *E. coli* O157:H7 prevalence (Fig. 1a). The first factor (41·0%) of PCA for the six OTUs separated the top 10% prevalence group from the bottom 10% prevalence group (Fig. S1).

In the MC diet group, 137 taxa and 790 OTUs were observed in 50% or more of the samples, and 2 taxa and 23 OTUs were correlated with *E. coli* O157:H7 prevalence (Tables 1 and 2). Negative correlations were detected ($P < 0{\cdot}05$) for order Bacillales ($r = -0{\cdot}21$) and genus *Butyrivibrio* ($r = -0{\cdot}17$) with *E. coli* O157:H7 prevalence (Table 1). However, no positive correlation was detected for taxa with *E. coli* O157:H7 prevalence.

Positive correlations were detected ($P < 0{\cdot}05$) for 13 of the 23 OTUs with *E. coli* O157:H7 prevalence. These 13 OTUs were assigned to S24-7 (two OTUs), Lachnospiraceae (one OTU), *Ruminococcus* (two OTUs), *Oscillospira* (one OTU), *Prevotella* (two OTUs), *Bacteroides* (one OTU) and *Prevotella copri* (four OTUs). On the other hand, negative correlations were detected ($P < 0{\cdot}05$) for the remaining 10 OTUs with *E. coli* O157:H7 prevalence. The 10 OTUs were assigned to Ruminococcaceae (one OTU), *Clostridium* (one OTU), *Dorea* (one OTU), *Roseburia* (one OTU), *Faecalibacterium prausnitzii* (one OTU), *Prevotella stercorea* (four OTUs) and *P. copri* (one OTU). Stepwise regression with backward elimination of 51 OTUs (OTUs with $P < 0{\cdot}1$, Table S1) showed that 15 OTUs each had significant impact on the model ($P < 0{\cdot}05$). Multiple linear regression for the 15 OTUs brought significant information to the model ($P < 0{\cdot}0001$) and explained 51·0% of variability in *E. coli* O157:H7 prevalence (Fig. 1b). The first factor (18·4%) of PCA for the 15 OTUs separated the top 10% prevalence group from the bottom 10% prevalence group (Fig. S1).

In the HC diet group, 146 taxa and 477 OTUs were observed in 50% or more of the samples, and 17 taxa and 24 OTUs were correlated with *E. coli* O157:H7 prevalence (Tables 1 and 2). Positive correlations were detected ($P < 0{\cdot}05$) for phylum Firmicutes ($r = 0{\cdot}19$), class Bacilli ($r = 0{\cdot}20$), order Turicibacterales ($r = 0{\cdot}19$) and Actinomycetales ($r = 0{\cdot}18$), family Turicibacteraceae ($r = 0{\cdot}19$),

Table 2 Commonly distributed OTUs observed in 50% or more of animals within each diet correlating (slope, nominal *P*-value and correlation coefficient) with *E. coli* O157:H7 prevalence in bovine faeces from diets with differing levels of concentrate

Diet*	#OTU ID	OTU taxonomy	Slope	*P*-value†	*R*
LC	None82399	p_Firmicutes; c_Clostridia; o_Clostridiales	0·37	0·001	0·34
	None327494	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Catabacteriaceae	0·31	0·030	0·22
	297920	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·29	0·007	0·28
	137650	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Ruminococcus	0·28	0·027	0·23
	None31867	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Butyrivibrio	0·28	0·024	0·23
	270019	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·33	0·042	0·21
	None87728	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·31	0·017	0·25
	None114996	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·29	0·031	0·22
	17822	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·35	0·036	0·22
	None53843	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·35	0·028	0·23
	None303364	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·32	0·042	0·21
	155911	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·31	0·042	0·21
	262668	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·31	0·029	0·23
	260468	p_Firmicutes; c_Erysipelotrichi; o_Erysipelotrichales; f_Erysipelotrichaceae	0·34	0·004	0·30
	262600	p_Firmicutes; c_Erysipelotrichi; o_Erysipelotrichales; f_Erysipelotrichaceae; g_Bulleidia	0·24	0·049	0·20
	None238285	p_Tenericutes; c_Mollicutes; o_RF39	0·35	0·042	0·21
	None212102	p_Tenericutes; c_Mollicutes; o_RF39	0·29	0·019	0·24
	None128757	p_Tenericutes; c_Mollicutes; o_RF39	0·25	0·033	0·22
	None100336	p_Tenericutes; c_Mollicutes; o_RF39	0·24	0·048	0·20
MC	None232659	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Paraprevotellaceae; g_Prevotella	0·12	0·048	0·17
	521417	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Bacteroidaceae; g_Bacteroides	0·24	0·018	0·20
	300859	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella	0·12	0·027	0·19
	329693	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	0·23	0·004	0·24
	524963	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	0·23	0·040	0·17
	None97371	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	0·16	0·049	0·17
	None340397	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	0·15	0·038	0·17
	177679	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_S24-7	0·31	0·017	0·20
	None237195	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_S24-7	0·18	0·020	0·20
	192225	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·14	0·023	0·20
	None92127	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Oscillospira	0·16	0·004	0·24
	269531	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·31	0·005	0·24
	None284176	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·20	0·008	0·22
	288932	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	−0·18	0·048	−0·17
	None182748	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_stercorea	−0·13	0·046	−0·17
	None361029	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_stercorea	−0·14	0·044	−0·17
	None156871	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_stercorea	−0·18	0·019	−0·20
	517331	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_stercorea	−0·25	0·002	−0·25
	287558	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Clostridiaceae; g_Clostridium	−0·17	0·006	−0·23
	293681	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Dorea	−0·24	0·008	−0·22
	168946	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Roseburia	−0·17	0·044	−0·17
	204099	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·18	0·047	−0·17
	173986	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Faecalibacterium; s_prausnitzii	−0·15	0·029	−0·18
HC	530894	p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_Corynebacteriaceae; g_Corynebacterium	0·21	0·008	0·23
	13354	p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_Corynebacteriaceae; g_Corynebacterium	0·20	0·003	0·26
	129049	p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_Propionibacteriaceae; g_Propionibacterium; s_acnes	0·15	0·027	0·19
	None73133	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·13	0·033	0·19
	20396	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·11	0·027	0·19
	341322	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·09	0·026	0·20
	None142341	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·13	0·015	0·21
	272886	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·09	0·025	0·20

Table 2 (*Continued*)

Diet*	#OTU ID	OTU taxonomy	Slope	*P*-value†	*R*
	None251969	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Ruminococcus	0·13	0·029	0·19
	77458	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Dorea	0·12	0·020	0·20
	41628	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·17	0·026	0·20
	68839	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·12	0·019	0·21
	None237036	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·11	0·042	0·18
	520434	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Veillonellaceae; g_Anaerovibrio	0·14	0·030	0·19
	521298	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_S24-7	−0·22	0·018	−0·21
	192071	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Dorea	−0·12	0·044	−0·18
	23757	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·16	0·023	−0·20
	308386	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·20	0·015	−0·21
	None97519	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	−0·11	0·031	−0·19
	521332	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	−0·12	0·043	−0·18
	176593	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	−0·13	0·009	−0·23
	528074	p_Firmicutes; c_Erysipelotrichi; o_Erysipelotrichales; f_Coprobacillaceae	−0·19	0·022	−0·20
	646411	p_Firmicutes; c_Erysipelotrichi; o_Erysipelotrichales; f_Erysipelotrichaceae; g_Bulleidia	−0·12	0·038	−0·18
	515507	p_Tenericutes; c_Mollicutes; o_RF39	−0·11	0·022	−0·20

*The diets were low concentrate (LC), moderate concentrate (MC) and high concentrate (HC). Widely distributed OTUs considered for each diet were 231, 790 and 477 OTUs respectively.

†The *P*-values presented represent nominal *P*-values used for the model generation. When corrected for false discovery using Bonferroni correction, none of the individual *P*-values reached a significance threshold $P < 0{\cdot}05$ for any of the diets.

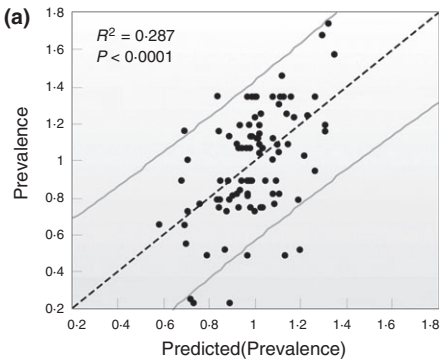
Corynebacteriaceae ($r = 0{\cdot}19$) and Propionibacteriaceae ($r = 0{\cdot}22$), genus *Corynebacterium* ($r = 0{\cdot}20$), *Turicibacter* ($r = 0{\cdot}19$) and *Propionibacterium* ($r = 0{\cdot}20$), and species *Propionibacterium acnes* ($r = 0{\cdot}20$) with *E. coli* O157:H7 prevalence (Table 1). On the other hand, negative correlations were detected ($P < 0{\cdot}05$) for phylum Cyanobacteira ($r = -0{\cdot}19$), class 4C0d-2 ($r = -0{\cdot}20$) and *Erysipelotrichi* ($r = -0{\cdot}21$), order YS2 ($r = -0{\cdot}21$) and Erysipelotrichales ($r = -0{\cdot}21$), and family Coprobacillaceae ($r = -0{\cdot}25$) with *E. coli* O157:H7 prevalence.

Positive correlations were detected ($P < 0{\cdot}05$) for 14 of the 24 OTUs with *E. coli* O157:H7 prevalence, whereas negative correlations were detected ($P < 0{\cdot}05$) for the remaining 10 OTUs with *E. coli* O157:H7 prevalence (Table 2). The positive 14 OTUs were classified to Lachnospiraceae (two OTUs), *Ruminococcus* (four OTUs), *Turicibacter* (three OTUs), *Corynebacterium* (two OTUs), *Dorea* (one OTU), *Anaerovibrio* (one OTU) and *Propionibacterium acnes* (one OTU). On the other hand, the negative 10 OTUs were classified to RF39 (one OTU), S24-7 (one OTU), Ruminococcaceae (two OTUs), Coprobacillaceae (one OTU), *Ruminococcus* (three OTUs), *Dorea* (one OTU) and *Bulleidia* (one OTU). Stepwise regression with backward elimination of 49 OTUs (OTUs with $P < 0{\cdot}1$, Table S1) showed that 10 OTUs had the significant impact on the model ($P < 0{\cdot}05$). Multiple linear regression for 10 OTUs brought significant information to the model ($P < 0{\cdot}0001$) and explained 33·1% of variability in *E. coli* O157:H7 prevalence (Fig. 1c). The first factor (24·6%) of PCA for the 10 OTUs separated the top 10% prevalence group from the bottom 10% prevalence group (Fig. S1).

The impact of faecal microbiome on *E. coli* O157:H7 enumeration

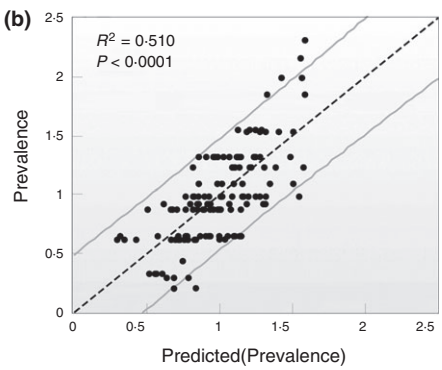
In the LC diet group, 8 taxa and 21 OTUs were correlated with *E. coli* O157:H7 enumeration (Tables 3 and 4) for the model building. Positive correlations were detected ($P < 0{\cdot}05$) for Enterobacteriales ($r = 0{\cdot}21$), Enterobacteriaceae ($r = 0{\cdot}21$), Coriobacteriales ($r = 0{\cdot}23$), Coriobacteriaceae $r = 0{\cdot}25$), Leuconostocaceae ($r = 0{\cdot}21$), *Leuconostoc* ($r = 0{\cdot}23$), *Lactococcus* ($r = 0{\cdot}25$) and *Anaerostipes* ($r = 0{\cdot}23$) with *E. coli* O157:H7 enumeration (Table 3). However, negative correlations were not detected for any taxon with *E. coli* O157:H7 enumeration.

Positive correlations were detected ($P < 0{\cdot}05$) for all 21 OTUs with *E. coli* O157:H7 enumeration (Table 4). The 21 OTUs were classified to CW040 (three OTUs), RF39 (one OTU), Clostridiales (one OTU), Ruminococcaceae (four OTUs), Erysipelotrichaceae (one OTU), Lachnospiraceae (one OTU), Clostridiaceae (one OTU), *Ruminococcus* (four OTUs), *Leuconostoc* (one OTU), *Turicibacter* (two OTUs), *Anaerostipes* (one OTU) and *Dorea* (one OTU). No OTUs with significant negative correlations were observed. Stepwise regression with backward elimination of 40 OTUs (OTUs with $P < 0{\cdot}1$, Table S2) indicated that three OTUs had significant impact on the model ($P < 0{\cdot}05$). Multiple linear regression for the three OTUs brought significant information to the model ($P < 0{\cdot}0001$) and explained



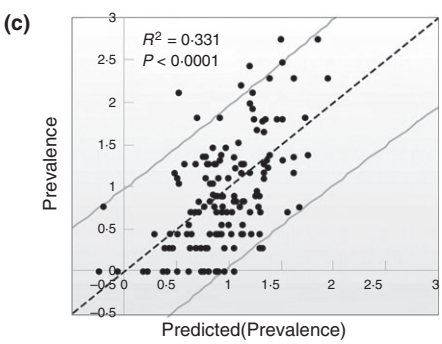
4 OTUs (positive correlation)
- order Clostridiales (1 OTU)
- family Ruminococcaceae (2 OTU)
- family Erysipelotrichaceae (1 OTU)

2 OTUs (negative correlation)
- family F16 (1 OTU)
- family Ruminococcaceae (1 OTU)



8 OTUs (positive correlation)
- family Lachnospiraceae (1 OTU)
- *Bacteroides (1 OTU)*
- *Oscillospira (1 OTU)*
- *Prevotella (1 OTU)*
- *Phascolarctobacterium (1 OTU)*
- *Ruminococcus (1 OTU)*
- *Prevotella copri (2 OTU)*

7 OTUs (negative correlation)
- class Clostridia (1 OTU)
- family Ruminococcaceae (1 OTU)
- *Clostridium (1 OTU)*
- *Coprococcus (1 OTU)*
- *Dorea (1 OTU)*
- *Prevotella (1 OTU)*
- *Prevotella copri (1 OTU)*



6 OTUs (positive correlation)
- family Lachnospiraceae (1 OTU)
- family Ruminococcaceae (1 OTU)
- *Anaerovibrio (1 OTU)*
- *Corynebacterium (1 OTU)*
- *Turicibacter (1 OTU)*
- *Propionibacterium acnes (1 OTU)*

4 OTUs (negative correlation)
- family Peptostreptococcaceae (1 OTU)
- family Ruminococcaceae (1 OTU)
- family S24 - 7 (1 OTU)
- *Bulleidia (1 OTU)*

Figure 1 Stepwise multiple linear regression with backward elimination for *Escherichia coli* O157:H7 prevalence across different diets. (a) Multiple linear regression for the six OTUs from faeces of cattle fed LC (low maize, forage-based) diet. Model explained 28·7% of variability (P < 0·0001). (b) Multiple linear regression for the 15 OTUs from faeces of cattle fed MC (moderate maize) diet. Model explained 51·0% of variability (P < 0·0001). (c) Multiple linear regression for the 10 OTUs from faeces of cattle fed HC (high maize) diet. Model explained 33·1% of variability (P < 0·0001).

22·8% of variability in *E. coli* O157:H7 enumeration (Fig. 2a). The first factor (75·6%) of PCA for the three OTUs separated the top 10% enumeration group from the bottom 10% enumeration group (Fig. S2).

In the MC diet, 4 taxa and 39 OTUs were correlated with *E. coli* O157:H7 enumeration (Table 4). Positive correlations were detected (P < 0·05) for Ruminococcaceae (r = 0·20) and *Ruminococcus bromii* (r = 0·18) with *E. coli* O157:H7 enumeration, whereas negative correlations were detected for *Escherichia* (r = −0·18) and putative genus rc4-4 (r = −0·18) with *E. coli* O157:H7 enumeration (Table 3).

Positive correlations were detected (P < 0·05) for 15 OTUs with *E. coli* O157:H7 enumeration (Table 4). The 15 OTUs were assigned to Bacteroidales (one OTU), RF39 (one OTU), Lachnospiraceae (two OTUs), Succinivibrionaceae (one OTU), S24-7 (one OTU), *Oscillospira* (one OTU), *Prevotella* (one OTU), *Ruminococcus* (three OTUs), *Succinivibrio* (one OTU), *Turicibacter* (one OTU) and *P. copri* (two OTUs). On the other hand, negative correlations were detected (P < 0·05) for 24 OTUs with *E. coli* O157:H7 enumeration. The 24 OTUs were assigned to Bacteroidales (four OTUs), Lachnospiraceae (two OTUs), S24-7 (two OTUs), *Phascolarctobacterium* (one OTU), *Succinivibrio* (one OTU), *Turicibacter* (one OTU), *Roseburia* (one OTU), *Prevotella* (eight OTUs), *P. copri* (one OTU) and *P. stercorea* (three OTU).

Stepwise multiple regression with backward elimination of 70 OTUs (OTUs with P < 0·1, Table S2) indicated that 12 OTUs had significant impact on the model (P < 0·05). Multiple linear regression for the 12 OTUs brought significant information on the model (P < 0·0001) and explained

Table 3 Commonly distributed taxa observed in 50% or more of animals within each diet correlating (slope, nominal *P*-value and correlation coefficient) with *Escherichia coli* O157:H7 enumeration in bovine faeces from diets with differing levels of concentrate

Diet*	Taxonomy	Slope	*P*-value†	*R*
LC	p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Leuconostocaceae	0·11	0·046	0·21
	p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Leuconostocaceae;g_Leuconostoc	0·11	0·028	0·23
	p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Lactococcus	0·11	0·016	0·25
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Anaerostipes	0·07	0·029	0·23
	p_Firmicutes;c_Clostridia;o_Coriobacteriales	0·05	0·029	0·23
	p_Firmicutes;c_Clostridia;o_Coriobacteriales;f_Coriobacteriaceae	0·06	0·016	0·25
	p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales	0·09	0·039	0·21
	p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae	0·09	0·039	0·21
MC	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae	0·02	0·018	0·20
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Ruminococcus;s_bromii	0·09	0·036	0·18
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Peptococcaceae;g_rc4-4	−0·08	0·035	−0·18
	p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae;g_Escherichia	−0·10	0·029	−0·18
HC	p_Firmicutes;c_Bacilli	0·05	0·015	0·21
	p_Firmicutes;c_Bacilli;o_Turicibacterales	0·06	0·039	0·18
	p_Firmicutes;c_Bacilli;o_Turicibacterales;f_Turicibacteraceae	0·06	0·039	0·18
	p_Firmicutes;c_Bacilli;o_Turicibacterales;f_Turicibacteraceae;g_Turicibacter	0·06	0·039	0·18
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Lachnobacterium	−0·07	0·008	−0·23
	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Roseburia;s_faecis	−0·12	0·018	−0·21
	p_Synergistetes	−0·07	0·045	−0·18
	p_Synergistetes;c_Synergistia	−0·07	0·045	−0·18
	p_Synergistetes;c_Synergistia;o_Synergistales	−0·07	0·045	−0·18

*The diets were low concentrate (LC), moderate concentrate (MC) and high concentrate (HC). Widely distributed taxa considered for each diet were 110, 137 and 146 taxa groups respectively.

†The *P*-values presented represent nominal *P*-values used for the model generation. When corrected for false discovery using Bonferroni correction, none of the individual *P*-values reached a significance threshold $P < 0·05$ for any of the diets.

38·8% of variability in *E. coli* O157:H7 enumeration (Fig. 2b). The first factor (20·6%) of PCA for the 12 OTUs separated the top 10% enumeration group from the bottom 10% enumeration group (Fig. S2).

In the HC diet, 9 taxa and 22 OTUs were correlated with *E. coli* O157:H7 enumeration (Tables 3 and 4). Positive correlations were detected for *Bacilli* ($r = 0·21$), Turicibacterales ($r = 0·18$), Turicibacteraceae ($r = 0·18$) and *Turicibacter* ($r = 0·18$) with *E. coli* O157:H7 enumeration (Table 3). On the other hand, negative correlations were detected for *Synergistetes* ($r = -0·18$), *Synergistia* ($r = -0·18$), Synergistales ($r = -0·18$), *Lachnobacterium* ($r = -0·23$) and *Roseburia faecis* ($r = -0·21$).

Positive correlations were detected ($P < 0·05$) for four OTUs with *E. coli* O157:H7 enumeration, and the four OTUs were assigned to Anaeroplasmataceae (one OTU), *Blautia* (one OTU) and *Turicibacter* (two OTUs) (Table 4). Negative correlations were detected ($P < 0·05$) for 18 OTUs with *E. coli* O157:H7 enumeration. The 18 OTUs were assigned to YS2 (one OTU), RF39 (one OTU), Lachnospiraceae (two OTUs), Ruminococcaceae (six OTUs), *Ruminococcus* (two OTUs), *Oscillospira* (one OTU), *R. faecis* (one OTU) and *F. prausnitzii* (four OTUs).

Stepwise multiple regression with backward elimination of 36 OTUs (OTUs with $P < 0·1$, Table S2) showed that 10 OTUs had significant impact on the model ($P < 0·05$). Multiple linear regression for the 10 OTUs brought significant information on the model ($P < 0·0001$) and explained 30·2% of variability in *E. coli* O157:H7 enumeration (Fig. 2c). The first factor (30·85%) of PCA for the 10 OTUs separated the top 10% enumeration group from the bottom 10% enumeration group (Fig. S2).

The impact of faecal bacterial diversity on *E. coli* O157:H7 prevalence and enumeration

Shannon's diversity index did not bring significant information to the simple linear regression model (data not shown). In addition, alpha diversity indices did not significantly differ between the top 10% prevalence (enumeration) and the bottom 10% prevalence (enumeration) groups except for the enumeration group in the MC diet group (Table S3). This result indicates that faecal bacterial diversity does not have an impact on *E. coli* O157:H7 prevalence and enumeration irrespective of diets.

Discussion

The composition of the bovine faecal microbiome can be highly variable, with diet having the strongest impact

on bacterial composition (Shanks *et al.* 2011; Kim *et al.* 2014). However, little is known regarding relationships between this diversity and animal phenotypes. In the current analyses conducted across multiple diets, taxa from 29 different phyla were observed. The HC diet was a high concentrate maize-based diet typical to finishing cattle in the United States, and we observed 26 different bacterial phyla from the composited faeces of the 132 animals fed this diet. In contrast, Xu *et al.* (2014) examined the bacterial microbiome in faeces from 22 finishing cattle in a feedlot study from Canada and reported taxa representing only 17 phyla. The current study provides the first extensive analysis of the faecal bacterial populations from large animal populations and the cattle-associated foodborne pathogen, *E. coli* O157:H7, whereas previous research by Xu *et al.* (2014) only examined 22 animals total. We attempt to make some comparisons to this previous research, but recognize that differences in collected pathogen phenotypes, animal populations, animal housing, animal diets, and bioinformatic procedures and analyses, makes direct comparison to this difficult.

Previous studies with intestinal pathogens have indicated that the infection and survival may be influenced by the bacterial diversity in the ecosystem. *Clostridium difficile* is a gastrointestinal pathogen that can cause diarrhoea in humans, and patients with recurring outbreaks of *C. difficile* were found to harbour a less diverse faecal microbiota that was highly variable between patients, than did normal patients (Chang *et al.* 2008). In agreement with this observation between diversity and a pathogen, lower microbial ecosystem diversity has been associated with longer survivability of *E. coli* O157:H7 in soils (Franz *et al.* 2008) and higher prevalence for *E. coli* O57:H7 and other enterohaemorrhagic *E. coli* on cattle hides (Chopyk *et al.* 2016). In contrast to the above studies, faecal microbiome studies with animals supershedding *E. coli* O157:H7 ($>10^4$ CFU per gram faeces) were found to have greater richness and diversity in the faecal microbiome than animals not shedding the pathogen (Xu *et al.* 2014). We have observed previously that the HC diet exhibited a high degree of variation in the faecal microbiota composition and taxa abundance across individual animals compared to the LC diet (Kim *et al.* 2014). As reported in our current study, the HC diet exhibited the lowest and the LC diet exhibited the highest prevalence and enumerable levels for *E. coli* O157:H7, contrary to the dogma that diversity drives pathogen opportunity. Furthermore, when we examined the pathogen phenotype extremes relationships to the alpha and beta diversities for the faecal samples, we did not observe any association between faecal bacterial diversity and *E. coli* O157:H7 phenotype within any of the three diets. Collectively, the data indicate that simple manipulation of the faecal microbiome diversity may not be a good approach to inhibit *E. coli* O157:H7 prevalence and enumeration in cattle.

Simple associations were examined between the microbiome taxa designations or OTU assignments and the phenotypic measures of *E. coli* O157:H7 prevalence and average enumerable levels. Overall for all diets, 44 individual taxa and 138 OTU groupings were found to have significant (nominal $P < 0{\cdot}05$) but weak linear associations with either *E. coli* O157:H7 phenotype, and only one of the OTU groupings (*Ruminococcus* sp. in the LC diet) was significant after stringent Bonferroni correction for false discovery. It was anticipated that correlations between *E. coli* O157:H7 phenotypes and the bovine faecal microbiome might be complex, and not limited to specific taxa groups. Furthermore, OTUs placed within taxa may correlate with an *E. coli* O157:H7 phenotype, whereas the larger phylogenetic taxa grouping may not correlate with the same *E. coli* O157:H7 phenotype. For example, *Prevotella* and *Ruminococcus* genera were not correlated with *E. coli* O157:H7 prevalence, but some OTUs placed within these genera were correlated with *E. coli* O157:H7 prevalence. This inconsistency indicates that different species from the same genus may exhibit different functions in faecal microbial ecosystems of cattle. Therefore, the use of OTUs rather than that of taxa will lead to better understanding of faecal microbiome correlations with *E. coli* O157:H7 phenotypes. In addition, the use of OTUs will be able to help understand correlations between unclassified taxa and *E. coli* O157:H7 because numerous sequences could not be assigned to a known genus against the current taxonomy.

In the current study with an average of 5660 sequence reads per sample, the majority of OTUs that correlated with *E. coli* O157:H7 prevalence and enumeration were not shared among the three diet groups. Faecal taxa or OTUs that correlated with *E. coli* O157:H7 prevalence and enumeration in cattle fed a diet may not be applicable to control *E. coli* O157:H7 prevalence and enumeration in cattle fed different diets. However, deeper sequencing will help determine whether OTUs observed in faeces of cattle fed one diet are much lower in abundance in animals fed a different diet. Nonetheless, based on the current research, separate experiments based on diets would need to be conducted to examine bovine faecal microbiomes that correlate with *E. coli* O157:H7 prevalence and enumeration in future studies.

Correlation relationships do not always represent causation, but the power of statistical correlations allows identification of associations that may be biologically related as levels of other microbiota may influence the ecosystem suitability for *E. coli* O157:H7. Bacterial species

Table 4 Commonly distributed OTUs observed in 50% or more of animals within each diet correlating (slope, nominal *P*-value and correlation coefficient) with *Escherichia coli* O157:H7 enumeration in bovine faeces from diets with differing levels of concentrate

Diet*	#OTU ID	OTU taxonomy	Slope	*P*-value†	*R*
LC	842413	p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Leuconostocaceae; g_Leuconostoc	0·17	0·005	0·29
	None140477	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·15	0·016	0·25
	20396	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·14	0·029	0·23
	None82399	p_Firmicutes; c_Clostridia; o_Clostridiales	0·08	0·024	0·23
	None347319	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Clostridiaceae	0·09	0·046	0·21
	None183121	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·13	0·012	0·26
	137650	**p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Ruminococcus**	**0·15**	**<0·0001**	**0·40**
	522551	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Ruminococcus	0·10	0·013	0·26
	99080	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Anaerostipes	0·14	0·024	0·23
	None235300	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Dorea	0·18	0·002	0·32
	None84753	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·14	0·001	0·33
	None320574	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·13	0·028	0·23
	None46278	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·11	0·034	0·22
	None65632	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	0·10	0·031	0·22
	None181956	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·16	0·003	0·30
	None347028	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·14	0·008	0·27
	260468	p_Firmicutes; c_Erysipelotrichi; o_Erysipelotrichales; f_Erysipelotrichaceae	0·12	0·001	0·33
	None68814	p_Tenericutes; c_Mollicutes; o_RF39	0·08	0·045	0·21
	None253223	p_TM7; c_TM7-3; o_CW040	0·15	0·003	0·30
	None333268	p_TM7; c_TM7-3; o_CW040	0·11	0·010	0·27
	None141672	p_TM7; c_TM7-3; o_CW040	0·13	0·021	0·24
MC	None210758	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales	0·06	0·048	0·17
	None116849	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella	0·07	0·032	0·18
	None44522	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	0·09	0·040	0·17
	None135639	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	0·08	0·027	0·19
	None88341	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_S24-7	0·09	0·044	0·17
	None334253	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·07	0·017	0·20
	526046	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·09	0·029	0·18
	None60398	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	0·09	0·018	0·20
	None92127	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Oscillospira	0·06	0·031	0·18
	527520	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·10	0·029	0·18
	None264115	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·09	0·045	0·17
	530805	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	0·08	0·007	0·22
	None264404	p_Proteobacteria; c_Gammaproteobacteria; o_Aeromonadales; f_Succinivibrionaceae	0·08	0·026	0·19
	None299047	p_Proteobacteria; c_Gammaproteobacteria; o_Aeromonadales; f_Succinivibrionaceae; g_Succinivibrio	0·06	0·028	0·19
	229874	p_Tenericutes; c_Mollicutes; o_RF39	0·08	0·028	0·18
	None25250	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales	−0·09	0·022	−0·19
	None278345	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales	−0·10	0·005	−0·23
	None25127	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales	−0·12	0·006	−0·23
	None58503	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales	−0·15	0·003	−0·25
	None205799	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Paraprevotellaceae; g_Prevotella	−0·08	0·025	−0·19
	None115050	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Paraprevotellaceae; g_Prevotella	−0·08	0·025	−0·19
	None131566	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Paraprevotellaceae; g_Prevotella	−0·08	0·045	−0·17
	20534	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Paraprevotellaceae; g_Prevotella	−0·09	0·037	−0·18
	None170592	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Paraprevotellaceae; g_Prevotella	−0·11	0·018	−0·20
	None164546	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella	−0·07	0·030	−0·18
	None224161	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella	−0·07	0·034	−0·18
	None258618	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella	−0·09	0·033	−0·18
	None15579	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_copri	−0·08	0·027	−0·19
	None129709	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella;s_stercorea	−0·08	0·013	−0·21
	None357586	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_stercorea	−0·08	0·033	−0·18
	None316901	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_Prevotellaceae; g_Prevotella; s_stercorea	−0·09	0·036	−0·18
	None297825	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_S24-7	−0·07	0·041	−0·17

Table 4 (*Continued*)

Diet*	#OTU ID	OTU taxonomy	Slope	*P*-value†	*R*
	None20751	p_Bacteroidetes; c_Bacteroidia; o_Bacteroidales; f_S24-7	−0·09	0·018	−0·20
	None294235	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	−0·09	0·039	−0·17
	None345423	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	−0·06	0·038	−0·18
	340642	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	−0·08	0·049	−0·17
	168946	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Roseburia	−0·09	0·040	−0·17
	516752	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Veillonellaceae; g_Phascolarctobacterium	−0·07	0·038	−0·18
	None173203	p_Proteobacteria; c_Gammaproteobacteria; o_Aeromonadales; f_Succinivibrionaceae; g_Succinivibrio	−0·09	0·005	−0·24
HC	None313813	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·09	0·031	0·19
	341322	p_Firmicutes; c_Bacilli; o_Turicibacterales; f_Turicibacteraceae; g_Turicibacter	0·06	0·048	0·17
	180067	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Blautia	0·09	0·041	0·18
	526376	p_Tenericutes; c_Mollicutes; o_Anaeroplasmatales; f_Anaeroplasmataceae	0·11	0·036	0·18
	515256	p_Cyanobacteria; c_4C0d-2; o_YS2	−0·11	0·027	−0·19
	None289779	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	−0·09	0·022	−0·20
	518613	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae	−0·10	0·045	−0·18
	510797	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Lachnospiraceae; g_Roseburia; s_faecis	−0·12	0·036	−0·18
	2835813	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·09	0·034	−0·19
	158273	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·10	0·019	−0·20
	23757	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·11	0·037	−0·18
	510377	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·11	0·006	−0·24
	None61296	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·11	0·013	−0·22
	17822	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae	−0·13	0·038	−0·18
	178357	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Faecalibacterium; s_prausnitzii	−0·08	0·024	−0·20
	187780	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Faecalibacterium; s_prausnitzii	−0·08	0·035	−0·18
	170981	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Faecalibacterium; s_prausnitzii	−0·09	0·039	−0·18
	520413	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Faecalibacterium; s_prausnitzii	−0·11	0·038	−0·18
	None360949	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Oscillospira	−0·08	0·033	−0·19
	511240	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	−0·11	0·020	−0·20
	None166118	p_Firmicutes; c_Clostridia; o_Clostridiales; f_Ruminococcaceae; g_Ruminococcus	−0·13	0·015	−0·21
	515507	p_Tenericutes; c_Mollicutes; o_RF39	−0·11	0·003	−0·26

*The diets were low concentrate (LC), moderate concentrate (MC) and high concentrate (HC). Widely distributed OTUs considered for each diet were 231, 790 and 477 OTUs respectively.

†The *P*-values presented represent nominal *P*-values used for the model generation. When corrected for false discovery using Bonferroni correction, none of the individual *P*-values reached a significance threshold $P < 0{\cdot}05$ for the MC or HC diets. One OTU did meet significance threshold for the LC diet, and this OTU is denoted in bold in the table.

corresponding to OTUs that were positively correlated with *E. coli* O157:H7 prevalence might provide nutrient sources for the growth of *E. coli* O157:H7. The LC diet was a forage-based diet and represents a type of diet fed to growing cattle, and in this dietary group faecal prevalence for *E. coli* O157:H7 averaged 66·4% positive and enumeration averaged 2·4 $\log_{10}$ CFU per gram faeces. The LC diet had 37 OTUs associated with either *E. coli* O157:H7 prevalence or enumeration, and all of these associations were positive. Three OTUs assigned to taxa were associated with both prevalence and enumeration. The majority of these OTUs assigned to taxa were placed within the order Clostridiales for both *E. coli* O157:H7 prevalence (68·4%) and enumeration (62·2%; Table S1). Because Clostridiales can degrade complex polysaccharides, including dietary fibre (Kamada *et al.* 2013), it might degrade ruminally undigested fibre into energy sources for the growth of *E. coli* O157:H7. An OTU assigned to the genus *Ruminococcus* had one of the highest observed correlation values ($R = 0{\cdot}4$) and a highly significant slope ($P < 0{\cdot}0001$). The OTUs assigned to RF39 placed within the class Mollicutes were positively correlated with both *E. coli* O157:H7 prevalence and enumeration. The genus *Mycoplasma* is assigned to Mollicutes, and species corresponding to these OTUs may be closely related to *Mycoplasma* that causes diseases such as

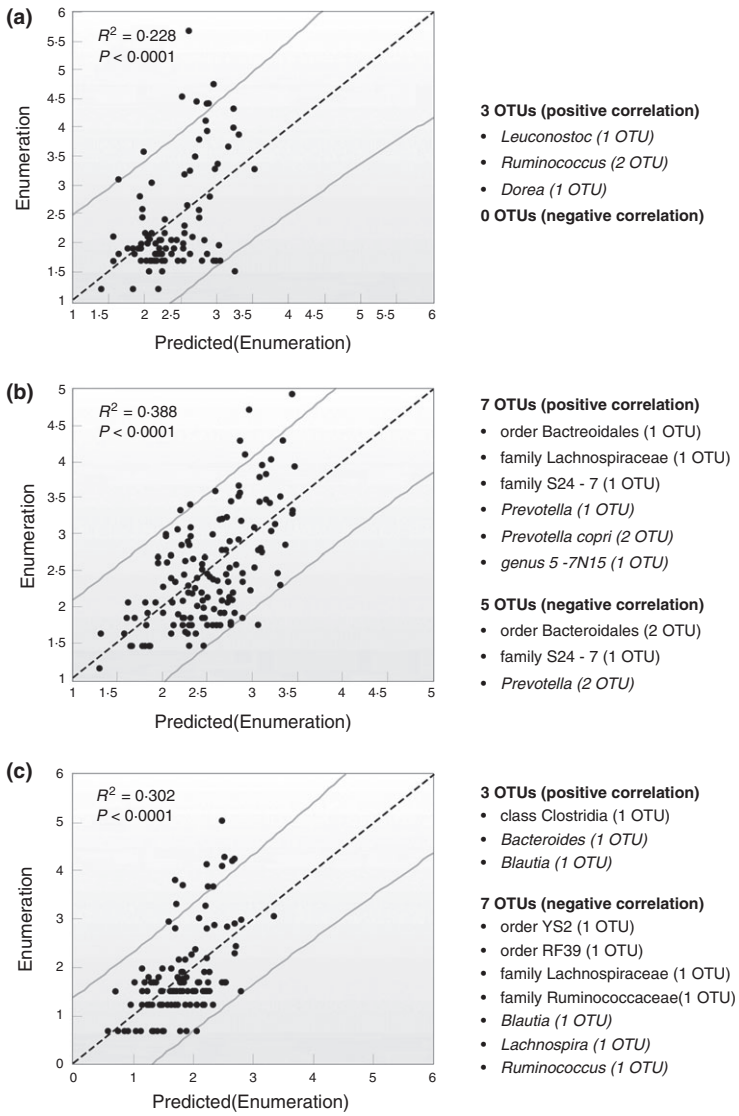


Figure 2 Stepwise multiple linear regression with backward elimination for *Escherichia coli* O157:H7 enumeration across different diets. (a) Multiple linear regression for the three OTUs from faeces of cattle fed LC (low maize, forage-based) diet. Model explained 22·8% of variability ($P < 0·0001$). (b) Multiple linear regression for the 12 OTUs from faeces of cattle fed MC (moderate maize) diet. Model explained 38·8% of variability ($P < 0·0001$). (c) Multiple linear regression for the 10 OTUs from faeces of cattle fed HC (high maize) diet. Model explained 30·2% of variability ($P < 0·0001$).

respiratory disease, otitis media, arthritis and mastitis in cattle (Maunsell *et al.* 2011). To date, no study has been conducted to examine an interaction between *Mycoplasma* and *E. coli* O157:H7, and further studies will need to be conducted to determine if there is a causative association between these bacterial groups.

The MC diet was a maize-based diet and represents a type of moderate-energy diet fed to late growing/early finishing cattle, and in this dietary group faecal prevalence for *E. coli* O157:H7 averaged 51·2% positive and enumeration averaged 2·5 $\log_{10}$ CFU per gram faeces. The MC diet had 60 OTUs associate with either *E. coli* O157:H7 prevalence or enumeration, and 26 of these associations were positive. Two of the OTUs assigned to taxa were associated with both prevalence and enumeration phenotypes. The majority of the OTUs assigned to *P. copri* in the MC diet group were positively correlated with *E. coli* O157:H7 prevalence and enumeration. Many *Prevotella* species are viewed as being beneficial, but in humans, *P. copri* may trigger rheumatoid arthritis and is associated with pro-inflammatory environments (Scher *et al.* 2013). Pathogenic bacteria can acquire a colonization and growth advantage from host inflammation that impedes commensal bacteria in the gastrointestinal tract (Kamada *et al.* 2013), thus bacteria like *P. copri* may play an inadvertent positive role in *E. coli* O157:H7 colonization. Similar to the LC diet, we observed OTUs assigned to the genus *Ruminococcus* that were positively associated with either *E. coli* O157:H7 prevalence or enumeration.

Concentrate diets like the HC diet are predominantly maize or another grain, and represent a typical

high-energy diet fed to finishing cattle. In this dietary group faecal prevalence for *E. coli* O157:H7 averaged 27·0% positive and enumeration averaged 1·8 $\log_{10}$ CFU per gram faeces. The HC diet had 43 of the OTUs associate with either *E. coli* O157:H7 prevalence or enumeration, and 17 of these associations were positive with one of the OTUs associated with both prevalence and enumeration phenotypes. The OTUs assigned to *Turicibacter* in the HC diet group were all positively correlated with both *E. coli* O157:H7 prevalence and enumeration. Although it has been reported that *Turicibacter* may be a pathogen in piglets (Cuiv *et al.* 2011), no study on *Turicibacter* relative to cattle health has been conducted. In a study of bovine feed efficiency, OTUs assigned to *Turicibacter* in the caecum were associated with differences in animal intake (Myer *et al.* 2015). *Turicibacter* appears to influence cattle performance but the causative interaction between *Turicibacter* and *E. coli* O157:H7 in cattle remains to be identified.

In contrast to those OTUs positively associated with *E. coli* O157:H7 prevalence and enumeration, bacterial species corresponding to OTUs that were negatively correlated with our *E. coli* O157:H7 phenotypes might produce bacteriocins inhibiting *E. coli* O157:H7, or these bacterial groups may compete with *E. coli* O157:H7 for nutrient sources or colonization sites. It should be noted that with the LC diet, no OTU was negatively associated with either *E. coli* O157:H7 phenotype. The lack of negative correlations and the observation that cattle fed the LC diet also had the highest *E. coli* O157:H7 prevalence needs to be investigated further.

In the MC diet group, 33 OTUs assigned to taxa were negatively associated with *E. coli* O157:H7, and nearly half of these OTUs were assigned to the genus *Prevotella*. Seven OTUs were assigned to *P. stercorea* that negatively correlated with *E. coli* O157:H7 prevalence or enumeration. *Prevotella* species seem to have a variety of functions in the faecal microbial ecosystem, but the potential role for strains of *P. stercorea* to influence *E. coli* O157:H7 is not apparent based on the limited literature for this species (Hayashi *et al.* 2007). Some OTUs assigned to taxa placed within Clostridiales also were negatively correlated with *E. coli* O157:H7 prevalence or enumeration. Species corresponding to these Clostridiales OTUs also seem to have a variety of functions in faecal microbial ecosystem.

In the HC diet group, 27 OTUs assigned to taxa were negatively associated with *E. coli* O157:H7, most of which were negatively associated with *E. coli* O157:H7 enumeration. The majority of OTUs that were negatively correlated with both *E. coli* O157:H7 prevalence and enumeration were assigned to taxa placed within Clostridiales, supporting that species corresponding to these Clostridiales OTUs play a wide range of roles in faecal microbial ecosystem. Previous research with cattle fed a high concentrate diet also recognized that taxa assigned within Clostridiales were differentially abundant in *E. coli* O157:H7 nonshedding and supershedding cattle (Xu *et al.* 2014). If isolated from bovine faeces and found to confer antagonism or inhibition to *E. coli* O157:H7, then the species corresponding to the OTUs that were negatively correlated with *E. coli* O157:H7 prevalence and enumeration may be useful as direct fed microbials to inhibit *E. coli* O157:H7.

Cumulative examination of the simple associations for all diets, nearly 58% of the OTUs that significantly associated with *E. coli* O157:H7 prevalence or enumeration, was positively correlated. These OTUs taxa assignments represented six phyla (Acitinobacter, Bacteroidetes, Firmicutes, Proteobacteria, Tenericutes and TM7) and ranged from *R*-values of 0·17–0·40. In contrast, the remaining assigned OTUs that were significantly associated negatively with our *E. coli* phenotypes were assigned to taxa of five phyla (Bacteroidetes, Cyanobacteria, Firmicutes, Proteobacteria and Tenericutes) and ranged from *R*-values of −0·17 to −0·26. For both the *E. coli* O157:H7 prevalence and enumeration associations, OTUs assigned to taxa belonged the Firmicutes were most represented (54·2 and 64·2% respectively). Bacteroidetes taxa groupings were the second most represented phyla, but more of the assigned OTUs were observed to negatively associate than positively associate (40·7 *vs* 17·3% respectively) with the *E. coli* O157:H7 phenotypes. The potential antagonism between species belonging to the phyla Bacteroidetes and *E. coli* O157:H7 needs to be further explored, and may be a target exploited through dietary manipulations.

Most of the individual OTUs listed above had potentially weak impact on *E. coli* O157:H7 prevalence and enumeration, however, it should be noted that bacteria do not always work alone. In contrast to previous research where simple comparisons of extreme groups are made, one advantage of a large study like we conducted is the ability to develop step-down regression models to estimate the contribution of groups of bacteria to variation in pathogen phenotypes. For stepwise multiple linear regression analyses, the criteria for inclusion were relaxed ($P < 0{\cdot}1$) to increase the OTU pool size for each diet and add additional dimensions to the statistical analyses. As observed through the stepwise multiple linear regression analyses, multiples OTUs together had much stronger associations on *E. coli* O157:H7 prevalence and enumeration than any single OTU. The results suggest, as expected, that the gastrointestinal microbial interactions that may drive *E. coli* O157 shedding are complex. Little information is available for many of the bacterial species observed in the multiple linear regression analyses. More

research needs to be conducted to better understand the bovine gastrointestinal ecosystems, particularly the microbial physiologies and genomics of relevant species, which would help identify specific enzymatic functions that could be exploited to alter the colonic microbiota. Furthermore, the use of a cocktail of species corresponding to multiple OTUs might be more effective than utilizing a single species corresponding to the observed OTUs to control *E. coli* O157:H7 prevalence and enumeration in cattle.

This study across multiple diets did not find common bacterial species that associated with *E. coli* O157:H7 prevalence and enumeration. Species corresponding to minor OTUs represented by small numbers of sequences also may be correlated with *E. coli* O157:H7 prevalence and enumeration. However, these minor OTUs were not analysed due to the lack of sequence reads. Deeper sequencing with high sequence reads will need to be conducted to examine the impact of the minor OTUs on *E. coli* O157:H7 prevalence and enumeration. Nonetheless, the current study provides information on species corresponding to predominant OTUs correlated with *E. coli* O157:H7 prevalence and enumeration. Metabolic reconstruction by metagenome sequencing (Albertsen *et al.* 2013) of faecal microbiome will lead to better understanding functions of the unknown species correlated with *E. coli* O157:H7 prevalence and enumeration and also may help conceive selective media to isolate these species. These isolated species may be useful as probiotics to inhibit *E. coli* O157:H7 prevalence and enumeration.

Acknowledgements

Mention of a trade name, proprietary product or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply approval to the exclusion of other products that may be suitable. USDA is an equal opportunity provider and employer. Dr. M. Kim was funded in part by a USDA/ARS Postdoctoral Award. Research was funded in part by USDA NIFA/AFRI Grant #2011-67005-30060. We appreciate the dedicated efforts provided by the USMARC Feedlot Crew for caring and handling of the cattle in this project, and are grateful for the excellent technical support of Dee Kucera, Shannon Ostdiek, Bruce Jasch, Sandy Fryda-Bradley and Steven Wahlmeier. In addition, we thank Jody Gallagher for exceptional secretarial assistance.

Conflict of Interest

The authors have not conflicts of interest to declare.

References

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W. and Nielsen, P.H. (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol* **31**, 533–538.

Arthur, T.M., Bosilevac, J.M., Brichta-Harhay, D.M., Guerini, M.N., Kalchayanand, N., Shackelford, S.D., Wheeler, T.L. and Koohmaraie, M. (2007) Transportation and lairage environment effects on prevalence, numbers, and diversity of *Escherichia coli* O157:H7 on hides and carcasses of beef cattle at processing. *J Food Prot* **70**, 280–286.

Arthur, T.M., Keen, J.E., Bosilevac, J.M., Brichta-Harhay, D.M., Kalchayanand, N., Shackelford, S.D., Wheeler, T.L., Nou, X. *et al.* (2009) Longitudinal study of *Escherichia coli* O157:H7 in a beef cattle feedlot and role of high-level shedders in hide contamination. *Appl Environ Microbiol* **75**, 6515–6523.

Berry, E.D. and Wells, J.E. (2010) *Escherichia coli* O157:H7: recent advances in research on occurrence, transmission, and control in cattle and the production environment. *Adv Food Nutr Res* **60**, 67–117.

Brichta-Harhay, D.M., Arthur, T.M., Bosilevac, J.M., Guerini, M.N., Kalchayanand, N. and Koohmaraie, M. (2007) Enumeration of *Salmonella* and *Escherichia coli* O157:H7 in ground beef, cattle carcass, hide and faecal samples using direct plating methods. *J Appl Microbiol* **103**, 1657–1668.

Callaway, T.R., Dowd, S.E., Edrington, T.S., Anderson, R.C., Krueger, N., Bauer, N., Kononoff, P.J. and Nisbet, D.J. (2010) Evaluation of bacterial diversity in the rumen and feces of cattle fed different levels of dried distillers grains plus solubles using bacterial tag-encoded FLX amplicon pyrosequencing. *J Anim Sci* **88**, 3977–3983.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L. and Knight, R. (2010a) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266–267.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G. *et al.* (2010b) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335–336.

Chang, J.Y., Antonopoulos, D.A., Kalra, A., Tonelli, A., Khalife, W.T., Schmidt, T.M. and Young, V.B. (2008) Decreased diversity of the fecal microbiome in recurrent *Clostridium difficile*-associated diarrhea. *J Infect Dis* **197**, 435–438.

Chopyk, J., Moore, R.M., DiSpirito, Z., Stromberg, Z.R., Lewis, G.L., Renter, D.G., Cernicchiaro, N., Moxley, R.A. *et al.* (2016) Presence of pathogenic *Escherichia coli* is correlated with bacterial community diversity and composition on pre-harvest cattle hides. *Microbiome* **4**, 9.

Cuiv, P.O., Klaassens, E.S., Durkin, A.S., Harkins, D.M., Foster, L., McCorrison, J., Torralba, M., Nelson, K.E. *et al.* (2011) Draft genome sequence of *Turicibacter sanguinis* PC909, isolated from human feces. *J Bacteriol* **193**, 1288–1289.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D. *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**, 5069–5072.

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G. and Edrington, T.S. (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol* **8**, 125.

Franz, E., Semenov, A.V., Termorshuizen, A.J., de Vos, O.J., Bokhorst, J.G. and van Bruggen, A.H. (2008) Manure-amended soil characteristics affecting the survival of *E. coli* O157:H7 in 36 Dutch soils. *Environ Microbiol* **10**, 313–327.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G., Ciulla, D., Tabbaa, D. *et al.*, Human Microbiome Consortium, Petrosino, J.F., Knight, R., Birren, B.W. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* **21**, 494–504.

Hayashi, H., Shibata, K., Sakamoto, M., Tomita, S. and Benno, Y. (2007) *Prevotella copri* sp. nov. and *Prevotella stercorea* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* **57**, 941–946.

Hu, Y., Zhang, Q. and Meitzler, J.C. (1999) Rapid and sensitive detection of *Escherichia coli* O157:H7 in bovine faeces by a multiplex PCR. *J Appl Microbiol* **87**, 867–876.

Kamada, N., Chen, G.Y., Inohara, N. and Nunez, G. (2013) Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol* **14**, 685–690.

Kim, M., Kim, J., Kuehn, L.A., Bono, J.L., Berry, E.D., Kalchayanand, N., Freetly, H.C., Benson, A.K. *et al.* (2014) Investigation of bacterial diversity in the feces of cattle fed different diets. *J Anim Sci* **92**, 683–694.

Low, J.C., McKendrick, I.J., McKechnie, C., Fenlon, D., Naylor, S.W., Currie, C., Smith, D.G., Allison, L. *et al.* (2005) Rectal carriage of enterohemorrhagic *Escherichia coli* O157 in slaughtered cattle. *Appl Environ Microbiol* **71**, 93–97.

Martinez, I., Kim, J., Duffy, P.R., Schlegel, V.L. and Walter, J. (2010) Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. *PLoS ONE* **5**, e15046.

Maunsell, F.P., Woolums, A.R., Francoz, D., Rosenbusch, R.F., Step, D.L., Wilson, D.J. and Janzen, E.D. (2011) *Mycoplasma bovis* infections in cattle. *J Vet Intern Med* **25**, 772–783.

Myer, P.R., Wells, J.E., Smith, T.P., Kuehn, L.A. and Freetly, H.C. (2015) Cecum microbial communities from steers differing in feed efficiency. *J Anim Sci* **93**, 5327–5340.

Nelson, M.C., Morrison, H.G., Benjamino, J., Grim, S.L. and Graf, J. (2014) Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. *PLoS ONE* **9**, e94249.

Neter, J., Wasserman, W. and Kutner, M.H. (1985) *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Models, 2nd Edition*, pp. 230–237. Homewood, IL: Irwin Publishing.

Omisakin, F., MacRae, M., Ogden, I.D. and Strachan, N.J. (2003) Concentration and prevalence of *Escherichia coli* O157 in cattle feces at slaughter. *Appl Environ Microbiol* **69**, 2444–2447.

Price, M.N., Dehal, P.S. and Arkin, A.P. (2010) FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490.

Rice, W.C., Galyean, M.L., Cox, S.B., Dowd, S.E. and Cole, N.A. (2012) Influence of wet distillers grains diets on beef cattle fecal bacterial community structure. *BMC Microbiol* **12**, 25.

Robinson, S.E., Wright, E.J., Hart, C.A., Bennett, M. and French, N.P. (2004) Intermittent and persistent shedding of *Escherichia coli* O157 in cohorts of naturally infected calves. *J Appl Microbiol* **97**, 1045–1053.

Scher, J.U., Sczesnak, A., Longman, R.S., Segata, N., Ubeda, C., Bielski, C., Rostron, T., Cerundolo, V. *et al.* (2013) Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife* **2**, e01202.

Shanks, O.C., Kelty, C.A., Archibeque, S., Jenkins, M., Newton, R.J., McLellan, S.L., Huse, S.M. and Sogin, M.L. (2011) Community structures of fecal bacteria in cattle from different animal feeding operations. *Appl Environ Microbiol* **77**, 2992–3001.

Wells, J.E., Shackelford, S.D., Berry, E.D., Kalchayanand, N., Bosilevac, J.M. and Wheeler, T.L. (2011) Impact of reducing the level of wet distillers grains fed to cattle prior to harvest on prevalence and levels of *Escherichia coli* O157:H7 in feces and on hides. *J Food Prot* **74**, 1611–1617.

Werner, J.J., Koren, O., Hugenholtz, P., DeSantis, T.Z., Walters, W.A., Caporaso, J.G., Angenent, L.T., Knight, R. *et al.* (2012) Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *ISME J* **6**, 94–103.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., McAllister, T.A. and Selinger, L.B. (2014) *Escherichia coli* O157:H7 super-shedder and non-shedder feedlot steers harbour distinct fecal bacterial communities. *PLoS ONE* **9**, e98115.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Principal component analysis for *E. coli* O157:H7 prevalence across different diets.

Figure S2. Principal component analysis for *E. coli* O157:H7 enumeration across different diets.

Table S1. OTUs associated with *E. coli* O157:H7 prevalence.

Table S2. OTUs associated with *E. coli* O157:H7 enumeration.

Table S3. Diversity measures for pathogen phenotypes.